

Figure 1A

NOVTRAN cDNA sequence:

atgcagtggtcctgtctggcctgtaccctcctcagggtcctcccacatgtcctgtctctcctgagagaccctgtgcctgtgcccacagggaccaagctcttc
cactcctgtatcacctcaacgaacccatgcgcctcctcctggagggtgctgttgaagctgcaggcatcacccctggactgtagggtctgagcaccgcc
ctgtccatatccatccctgcctgctcctcctgacccgactccttcaacagaccatccctgctccttcaacaggccccgctctgctggggaacca
cggacagaggccttcccatcccaggcctgaaggccagagtaggtggaccatcctgccgaagccggcctcaattctcaaggccatgccgtggagc
cagtgccatctggaccctctgggtcaagcaaaagggtgtgtgctaatcaaaaggcagccctcgaggatgccaaaggccccgaatgccagtggaccg
tgaaaaccttctgtgacaaaacctgcagtgccttctctgtcctcagctgctctccagctcctcatcaagggtgaaacagagcaggagcgcagtaatg
cggaatttgacttgcaaaagtcgggcccgtcgggattacaattcaaggctgctgctgaaactcgggcagatcccagctgcaaaaggcagttccttctcga
gctgcagaacgtgtctggagggttggtcagcccaggtcccaggaaacctgcaagggtggggcgggccctcagagcccttccagagctggg
ggctggtagccccctttggctttggagaaggtcagtaaccaaccttccccaggcccgcactgcggaagggtgtggactggccccctgtgtcctggt
gaccagtgtccactgtgcacttcccaggccagccgaacctggcacacactgggtgttcctaaatagccatggagggtattgtggcatggagagctgtc
gattccagaaacctcctggacatagggtggtggagctcatctgcagaagctgcctga

Figure 1B

NOVTRAN Protein Sequence:

MQWSCLACTLLRVLPHVLSLLRDPVPVPTGTKLFHSCITSTNPCASFLEVAVEAAGITPW
TVGSEHPPCPYP SLHASPF TDSFN RSPAPLNRPRSAGEPRTEAFSPGLKARVGGTILAE
AGLNSQGHAVEPVPSGPSGSSKGCVLIKGRPSRMPKARECPVDRENLLL TNPAVPSLLQL
LSSSPCIKVETE QERSNAEFDLQ SRAARDYNSRLLLKL GQIPAAKGSSFLELQNVSGGVG
SARGPRNHCKV GAGPQSPFELGAGSPPLALEKVSTQPIQARLRKGVDWPPVSPGDQC
PLCTLPGQPNLAHTGCSLNSHGGYCGMESCRFQKPPGHRAGSSSAEAA

gb:GENBANK-ID:H55724|acc:H55724 CHR220663 Chromosome 22 exon Homo sapiens
genomic clone C22 916 5', mRNA sequence - Homo sapiens, 171 bp.

Score = 585 (87.8 bits), Expect = 1.0e-19, P = 1.0e-19
Identities = 117/117 (100%), Positives = 117/117 (100%), Strand = Minus / Plus

```

Query:   735  ACCTCGGGCTGAGCCAACCCCTCCAGACACGTTCTGCAGCTCGAGGAAGGAAGTGCCTT  676
          |||
Sbjct:   18   ACCTCGGGCTGAGCCAACCCCTCCAGACACGTTCTGCAGCTCGAGGAAGGAAGTGCCTT  77

Query:   675  TGCAGCTGGGATCTGCCCAGTTTCAGCAGCAGCCTTGAATTGTAATCCCAGCGGC  619
          |||
Sbjct:   78   TGCAGCTGGGATCTGCCCAGTTTCAGCAGCAGCCTTGAATTGTAATCCCAGCGGC  134

```

[illegible]

Figure 2B

BlastX for NOVTRAN:

No Hits:

ptnr:SWISSPROT-ACC:P22357 ANOTHER-SPECIFIC PROTEIN SF18...	-3	102	0.0093	1
ptnr:SPTREMBL-ACC:Q26876 PROTECTIVE RECOMBINANT ANTIGE...	-3	87	0.018	1
ptnr:TREMBLNEW-ACC:AAD41775 PRO-ALPHA-2(I) COLLAGEN - ...	-3	86	0.061	2
ptnr:PIR-ID:S46272 anther-specific protein - common su...	-3	94	0.073	1
ptnr:SPTREMBL-ACC:Q26874 PROTECTIVE RECOMBINANT ANTIGE...	-3	78	0.16	1
ptnr:SPTREMBL-ACC:Q92029 ALPHA-1 COLLAGEN TYPE III - G...	-3	86	0.19	3
ptnr:SWISSPROT-ACC:P46804 SPIDROIN 2 (DRAGLINE SILK FI...	-3	101	0.24	1
ptnr:SPTREMBL-ACC:O57808 103AA LONG HYPOTHETICAL PROTE...	+2	76	0.25	1
ptnr:SWISSNEW-ACC:P41484 PROLINE-RICH ANTIGEN (36 KD A...	-3	95	0.26	1

CuraBLASTX Analysis of AC007663_A

PUBLIC DATABASE

Query= AC007663_A Cura_109 transmembrane protein
(1047 letters)

Database: /opt/database/public/blast/protnr
577,633 sequences; 178,813,065 total letters.

			Smallest Sum	Reading Frame	High Score	Probability P(N)	N
Sequences producing High-scoring Segment Pairs:							
ptnr:SWISSPROT-ACC:P22357 ANOTHER-SPECIFIC PROTEIN SF18...	-3	102	0.015	1			
ptnr:pir-id:S12246 anther-specific protein SF18 precur...	-3	102	0.015	1			

PATENT DATABASE

CuraBLASTX Analysis of AC007663_A

Query= AC007663_A Cura_109 transmembrane protein
(1047 letters) of query sequence in all 6 reading frames

Database: /opt/database/licensed/blast/geneseq_aa
349,121 sequences; 51,277,408 total letters.

			Smallest Sum	Reading Frame	High Score	Probability P(N)	N
Sequences producing High-scoring Segment Pairs:							
patp:R28916 Type III procollagen (prior art) - Homo sa...	+1	81	0.059	3			
patp:W03454 Mutant mouse c-AKT non-translation region ...	+1	81	0.075	1			
patp:G04146 Arabidopsis thaliana protein fragment SEQ ...	-3	80	0.095	1			

Figure 3A

NOVNEUR cDNA sequence:

Regions 135319 to 135129 (1-190bp) and 134487 to 134314 (191 to 364bp) of the genomic clone (in - frame). Further extended using ESTs AI739078, AI767488, and AI796217.

AGCGCGCCCGAACGAAGCCGCGGCCCGGGCACAGCATGGCCCGCGGCGGGAGGGCGCTCGGATGTTTCGGC
AGCCTCCTGCACTTCGCCCCTGCTCGCTGCCGGCGTCTGTCCTCCGCTCAGCTGGGATCTCCCGAGCCCCGCA
GCCGAGCCAGCAAGATCCGAGTGCCTCGCGAGGCAAGCTCTGGGCCATCGGTCACTTCATGGGCAAGAA
GAGTCTGGAGCCTTCCAGCCCATCCCCATTGGGGACAGCTCCCCACACCTCCCTGAGGGACCAGCGACTG
CAGCTGAGTCATGATCTGCTCGGAATCCTCCTGCTAAAGAAGGCTCTGGGCGTGAGCCTCAGCCGCCCCG
CACCCCAAATCCAGTACAGGAGGCTGCTGGTACAAATACTGCAGAAATGACACCAATAATGGGGCAGACA
CAACAGCGTGGCTTAGATTGTGCCCCACCCAGGGAAGGTGCTGAATGGGACCCTGTTGATGGCCCCATCTG
GATGTAAATCCTGAGCTCAAATCTCTGTTACTCCATTACTGTGATTTCTGGCTGGGTCAACAGAAATATC
GCTGATGCAGACACAGATTATGTTCTGCTGTATTTCTGCTTCCCTGTTGAATTGGTGAATAAAACCTT
GCTCTATACATACAAA

Figure 3B

NOVNEUR Protein sequence:

MFGSLLHFALLAAGVVPLSWDLPEPRSRASKIRVHSRGLWAIGHFMGKKSLEPSSPSPLGTAPHTSLRDQRLQLSH
DLLGILLKKALGVSLRPAPQIQYRLLVQILQK

Figure 4A

BlastN for NOVNEUR:

gb:GENBANK-ID:HUMNKB|acc:M21551 Human neuromedin B mRNA, complete cds -
Homo sapiens, 640 bp (RNA).

Top Previous Match Next Match
Length = 640

Plus Strand HSPs:

Score = 3055 (458.4 bits), Expect = 9.1e-132, P = 9.1e-132
Identities = 629/639 (98%), Positives = 629/639 (98%), Strand = Plus / Plus

```

Query:      2 GCGCGCCCGAACGAAGCCGCGGCCCGGGCACAGC-ATGGCCCG-CGG-CGGGAGGGCGCT 58
             GCGCGCCCGAACGAAGCCGCGGCCCGGGCACAGC ATGGCCCG CGG CGGG GGGCGCT
Sbjct:      2 GCGCGCCCGAACGAAGCCGCGGCCCGGGCACAGCCATGGCCCGCGGGCGGG-GGGCGCT 60

Query:     59 CGGATGTTTCGGCAGCCTCCTGCACTTCGCCCTGCTCGCTGCCGGCGTCTGTCCTCCGCTCAGC 118
             CGGATGTTTCGGCAGCCTCCTGCA CTTGCCCTGCTCGCTGCCGGCGTCTG CCGCTCAGC
Sbjct:     61 CGGATGTTTCGGCAGCCTCCTGCTCTTCGCCCTGCTCGCTGCCGGCGTCTGCCCCGCTCAGC 120

Query:    119 TGGGATCTCCCGGAGCCCCGCGAGCCGAGCCAGCAAGATCCGAGTGCACTCGCGAGGCAAG 178
             TGGGATCTCCCGGAGCCCCGCGAGCCGAGCCAGCAAGATCCGAGTGCACTCGCGAGGCAA
Sbjct:    121 TGGGATCTCCCGGAGCCCCGCGAGCCGAGCCAGCAAGATCCGAGTGCACTCGCGAGGCAAC 180

Query:    179 CTCTGGGGCCATCGGTCACTTCATGGGCAAGAAGAGTCTGGAGCCTTCCAGCCCATCCCCA 238
             CTCTGGGGCCA CGGTCACTTCATGGGCAAGAAGAGTCTGGAGCCTTCCAGCCCATCCC A
Sbjct:    181 CTCTGGGGCCACCGTCACTTCATGGGCAAGAAGAGTCTGGAGCCTTCCAGCCCATCCC-A 239

Query:    239 TTGGGGACAGCTCCCCACACCTCCC-TGAGGGACCAGCGACTGCAGCTGAGTCATGATCT 297
             TTGGGGACAGCTCCCCACACCTCCC TGAGGGACCAGCGACTGCAGCTGAGTCATGATCT
Sbjct:    240 TTGGGGACAGCTCCCCACACCTCCCCCTGAGGGACCAGCGACTGCAGCTGAGTCATGATCT 299

Query:    298 GCTCGGAATCCTCCTGCTAAAGAAGGCTCTGGGCGTGAGCCTCAGCCGCCCCGCAACCCCA 357
             GCTCGGAATCCTCCTGCTAAAGAAGGCTCTGGGCGTGAGCCTCAGCCGCCCCGCAACCCCA
Sbjct:    300 GCTCGGAATCCTCCTGCTAAAGAAGGCTCTGGGCGTGAGCCTCAGCCGCCCCGCAACCCCA 359

Query:    358 AATCCAGTACAGGAGGCTGCTGGTACAAATACTGCAGAAATGACACCAATAAT-GGGGCA 416
             AATCCAGTACAGGAGGCTGCTGGTACAAATACTGCAGAAATGACACCAATAAT GGGGCA
Sbjct:    360 AATCCAGTACAGGAGGCTGCTGGTACAAATACTGCAGAAATGACACCAATAATAGGGGCA 419

Query:    417 GACACAACAGCGTGGCTTAGATTGTGCCCACCCAGGGAAGGTGCTGAATGGGACCCTGTT 476
             GACACAACAGCGTGGCTTAGATTGTGCCCACCCAGGGAAGGTGCTGAATGGGACCCTGTT
Sbjct:    420 GACACAACAGCGTGGCTTAGATTGTGCCCACCCAGGGAAGGTGCTGAATGGGACCCTGTT 479

Query:    477 GATGGCCCCATCTGGATGTAAATCCTGAGCTCAAATCTCTGTTACTCCATTACTGTGATT 536
             GATGGCCCCATCTGGATGTAAATCCTGAGCTCAAATCTCTGTTACTCCATTACTGTGATT
Sbjct:    480 GATGGCCCCATCTGGATGTAAATCCTGAGCTCAAATCTCTGTTACTCCATTACTGTGATT 539

Query:    537 TCTGGCTGGGTCAACAGAAATATCGCTGATGCAGACACAGATTATGTTTCTGCTGTATTT 596
             TCTGGCTGGGTCAACAGAAATATCGCTGATGCAGACACAGATTATGTTTCTGCTGTATTT
Sbjct:    540 TCTGGCTGGGTCAACAGAAATATCGCTGATGCAGACACAGATTATGTTTCTGCTGTATTT 599

Query:    597 CCTGCTTCCCTGTTGAATTGGTGAATAAAACCTTGCTCTAT 637
             CCTGCTTCCCTGTTGAATTGGTGAATAAAACCTTGCTCT T
Sbjct:    600 CCTGCTTCCCTGTTGAATTGGTGAATAAAACCTTGCTCTTT 640

```

Figure 4B

BlastX for NOVNEUR:

ptnr:SWISSPROT-ACC:P08949 NEUROMEDIN B-32 PRECURSOR [CONTAINS: NEUROMEDIN
B] - Homo sapiens (Human), 121 aa.

Top Previous Match Next Match

Length = 121

Plus Strand HSPs:

Score = 513 (180.6 bits), Expect = 1.9e-48, P = 1.9e-48

Identities = 105/118 (88%), Positives = 105/118 (88%), Frame = +2

Query: 44 RREGARMFGSLLHFALLAAGVVPLSWDLPEPRSRASKIRVHSRGKLWAIGHFMGKKSLEP 223
R GARMFGSLL FALLAAGV PLSWDLPEPRSRASKIRVHSRG LWA GHFMGKKSLEP

Sbjct: 4 RAGGARMFGSLLLLFALLAAGVAPLSWDLPEPRSRASKIRVHSRGNLWATGHFMGKKSLEP 63

Query: 224 SSPSPLGTAPHTSLRDQRLQLSHDLLGILLKKALGVSLSRPAPQIQYRLLVQILQK 397
SSPS G P LRDQRLQLSHDLLGILLKKALGVSLSRPAPQIQYRLLVQILQK

Sbjct: 64 SSPSHWGQLPTPLRDQRLQLSHDLLGILLKKALGVSLSRPAPQIQYRLLVQILQK 121

Query: 44 RREGARMFGSLLHFALLAAGVVPLSWDLPEPRSRASKIRVHSRGKLWAIGHFMGKKSLEP 223
R GARMFGSLL FALLAAGV PLSWDLPEPRSRASKIRVHSRG LWA GHFMGKKSLEP
Sbjct: 4 RAGGARMFGSLLLLFALLAAGVAPLSWDLPEPRSRASKIRVHSRGNLWATGHFMGKKSLEP 63
Query: 224 SSPSPLGTAPHTSLRDQRLQLSHDLLGILLKKALGVSLSRPAPQIQYRLLVQILQK 397
SSPS G P LRDQRLQLSHDLLGILLKKALGVSLSRPAPQIQYRLLVQILQK
Sbjct: 64 SSPSHWGQLPTPLRDQRLQLSHDLLGILLKKALGVSLSRPAPQIQYRLLVQILQK 121

Figure 5

ClustalW for NOVNEUR:

Based on this alignment, black outlined amino acid residues indicate regions of conserved sequence (i.e., regions that may be required to preserve structural or functional properties); greyed amino acid residues can be mutated to a residue with comparable steric and/or chemical properties without altering protein structure or function (e.g., L to V, I, or M); non-highlighted amino acid residues can potentially be mutated to a much broader extent without altering structure or function.

Sequences used:

- 1) A37178 (neuromedin B precursor – rat): Locus A37178, accession A37178, PID - g112182
- 2) NEUB_HUMAN (NEUROMEDIN B-32 PRECURSOR): Locus - NEUB_HUMAN, accession - P08949, PID - g1346684

```
Neuromedin_New_REVCOMP  -----MFGSLLHFALLAAGVPLSWDLPEPRSRASKIRVHSRGKLWALGHFMGKKS
A37178                  MTRQAGSTWMLLRGLLLFALLFVSGITPFSWDLPEPRSRASKIRVHPRGNLWATGHFMGKKS
NEUB_HUMAN              MARRAGGARMFGSLLLFAALLAAGVAPLSWDLPEPRSRASKIRVHSRGNLWATGHFMGKKS

Neuromedin_New_REVCOMP  LEPSSPSPGLGTAPHTSLRDQRLQLSHDLLGILLKKALGVSLSRPAPQIQYRLLVQILQ
A37178                  LEPSSLSLVGTAPPIITQREQRLQLSHDLLRIILLQKALGMNLSGPAPPIQYRLLQK---
NEUB_HUMAN              LEPSSPSHWGQLPTPPLRDQRLQLSHDLLGILLKKALGVSLSRPAPQIQYRLLVQILQ

Neuromedin_New_REVCOMP  K
A37178                  -
NEUB_HUMAN              K
```

Figure 6A

NOVGON cDNA sequence:

1
ATGAAGCTGGCATTCTCTTCTTGGCCCCATGGCCCTCCTCCTTCTGGCTGGCTATGGCTGTGTCCTCGGTGCCTCCAG
81
TGGGAACCTGCGCACCTTTGTGGGCTGTGCCGTGAGGGAGTTTACTTTCCTGGCCAAGAAGCCAGGCTGCAGGGGCCTTC
161
GGATCACCACGGATGCCTGCTGGGGTCGCTGTGAGACCTGGGAGAAACCCATTCTGGAACCCCCCTATATTGAAGCCCAT
241
CATCGAGTCTGTACCTACAACGAGACCAAACAGGTGACTGTCAAGCTGCCCAACTGTGCCCCGGGAGTCGACCCCTTCTA
321
CACCTATCCCGTGGCCATCCGCTGTGACTGCCGAGCCTGCTCCACTGCCACCACGGAGCTGAGGTTGATGCCAGGGGAAG
401
CTGCTGTGGCAC'TGGGCTTCTGGTGT'CAGCGTAGGAGACAGGGATCTAGGACAACAGGGACCAGGTGGCGACATGCAGCT
481
GTAAGAGACAAGGTGAGTCTCCTGAAGGCAGTAGATGGTTGGAATGGGCTGCTTGGGGACCCAGCGAGCTCCCAGGGCCT
561
TTCTGCTTCTTCTGTACCCCTGTATTTCCCTTGGCTTTCCAAATTGACTCAGCTTCTGGTAAAGTTGGAAACTTTTCCA
641
GCAAACAGACCTTCATCTTCTCCAGTGCAGAGATTACATTAGGAGGAACATGA

Figure 6B

NOVGON Protein Sequence:

MKLAFLFLGPMALLLLAGYGCVLGASSGNLRTFVGCAVREFTFLAKKPGCRGLRITTD
CWGRCEWKEPILEPPYIEAHRVCTYNETKQVTVKLPNCAPGVDPFYTYPVAIRCDG
ACSTATTELRLMPGEAAVALGFWCQRRRQGSRTTGTRWRHA AVRDKVSLKAVD
GW
NLLGDPASSQGLSASSCTPVFPLAFQIDSASGKVGNFSSKQTFIFSSAEITLGGT*

Figure 7A

BlastN for NOVGO:

AF146151 Salmo salar gonadotropin II beta subunit mRNA,
partial cds - Salmo salar, 266 bp (RNA).

Top Previous Match Next Match
Length = 266

Plus Strand HSPs:

Score = 208 (31.2 bits), Expect = 0.043, P = 0.042

Identities = 124/205 (60%), Positives = 124/205 (60%), Strand = Plus / Plus

Query: 177 CTGCTGGGGTCGCTGTGAGACCTGGGAGAAACCCATTCTGGAACCCCCCTATATTGAAGC 236
CTGC G GG C CTG G ACC GGAG CC TT T AA CCC AT TT A C
Sbjct: 34 CTGCAGTGGCCACTGCGTCACCAAGGAG---CCGGTTTTTC-AAGAGCCC-ATTTTCCACC 88

Query: 237 CCAT-C-ATCGAGTCTGTACCTACAACGA-GACCAAACAGGTGACTGTCAAGCTGCCCAA 293
T C A C GT TG ACCTAC GA G CC A G AC TC CT CC A
Sbjct: 89 GTGTACCAGCATGTGTGCACCTACCGGGACGTCCGCT-ATGAAACGATCCGCCTACCTGA 147

Query: 294 CTGTGCCCCGGGAGTCGACCC-CTTCTACACCTATCCCGTGGCCATCCGCTGTGACTGCG 352
CTGT CCCC G GT GACC C T T CACCTA CC GTGGC T GCTGTGACTGC
Sbjct: 148 CTGTCCCCCTTGGGTGGACCATCATGT-CACCTACCCTGTGGCTCTGAGCTGTGACTGCA 206

Query: 353 GAGCCTGCTCCACTGCCACCACGGAGCTG 381
G CTG CA G CAC C GA CTG
Sbjct: 207 GCCTCTGTAACATGGACACTTCTGA-CTG 234

Figure 7B

BlastX for NOVGON:

P01235 GONADOTROPIN BETA CHAIN PRECURSOR - *Cyprinus carpio* (Common carp), 144 aa.

Score = 203 (71.5 bits), Expect = 2.3e-16, P = 2.3e-16
Identities = 37/85 (43%), Positives = 52/85 (61%)

Query: 42 TFLAKKPGC-RGLRITTDACWGRCEWKEPPIEYIAHHRVCTYNETKQVTVKLPNCA 100
T +K GC + L + T C G C T E P+ + P+ + VCTY + + TV+LP+C
Sbjct: 39 TVAVEKEGCPKCLVLQTTICSGHCLTKE-PVYKSPFSTVYQHVCTYRDVRYETVRLPDCP 97

Query: 101 PGVDPFYTYTPVAIRCDGACSTATTE 126
PGVDP TYPVA+ CDC C+ T++
Sbjct: 98 PGVDPHITYPVALSCDCSLCTMDTSD 123

>ptnr:SWISSNEW-ACC:Q9YGH2 GONADOTROPIN BETA-II CHAIN PRECURSOR (GTH-II-BETA)
- *Clupea pallasii* (Pacific herring), 149 aa.

Plus Strand HSPs:

Score = 206 (72.5 bits), Expect = 1.0e-15, P = 1.0e-15
Identities = 47/117 (40%), Positives = 63/117 (53%), Frame = +1

Query: 28 PMALLLLAGYGCVLG--ASSGNLRTFVGCAVREFTFLAKKPGC-RGLRITTDACWGRCEW 198
P +LL CVL A NL+ C + T +K GC R L T C G C T
Sbjct: 5 PECTILLLLCMCVLAVPAQCFNLQP---CVLVNETVSVEKEGCPRLVFRTTICSGHCPT 61

Query: 199 WEKPILEPPYIEAHHRVCTYNETKQVTVKLPNCAPGVDPFYTYTPVAIRCDGACSTATTE 378
E P+ + P+ + VCTY + T++LP+CA GVDP TYPVA+ C+C CS T++
Sbjct: 62 KE-PVYKSPFSVNVQHVCTYGNFRYETIRLPDCADGVDPPLVTYPVALSCECSLCSMDTSD 120

>patp:R15106 hCG/bLH chimera, D10 - *Homo sapiens*, 145 aa.

Plus Strand HSPs:

Score = 188 (66.2 bits), Expect = 5.0e-14, P = 5.0e-14
Identities = 40/101 (39%), Positives = 58/101 (57%), Frame = +1

Query: 76 SSGNLRTFVGCAVREFTFLAKKPGCRG-LRITTDACWGRCEWKE--PILEPPYIEAHR 246
S G LR C T A+K C + TT C G C + ++ P++ PP + R
Sbjct: 1 SRGPLRPL--CQPINATLAAEKEACPVCITFTTSICAGYCPSMKRVLPVILPPMPQ---R 55

Query: 247 VCTYNETKQVTVKLPNCAPGVDPFYTYTPVAIRCDGACSTATTE 378
VCTY+E + +V+LP C PGVDP ++PVA+ C CG C ++T+
Sbjct: 56 VCTYHELRFASVRLPGCPPGVDPMVSPVALSCHGPCRLSSTD 99

>patp:W65110 Equine chorionic gonadotropin beta-chain protein - *Equus caballus*, 169 aa.

Plus Strand HSPs:

Score = 175 (61.6 bits), Expect = 1.2e-12, P = 1.2e-12
Identities = 43/116 (37%), Positives = 60/116 (51%), Frame = +1

Query: 31 MALLLLAGYGCVLGASSGNLRTFVGCAVREFTFLAKKPGCR-GLRITTDACWGRCEWKE 207
+ L +L G V AS G LR C T A+K C + TT C G C + +
Sbjct: 7 LLLWMLLSVGGVW-ASRGPLRPL--CRPINATLAAEKEACPICITFTTSICAGYCPSMVR 63

Query: 208 PILEPPYIEAHR-VCTYNETKQVTVKLPNCAPGVDPFYTYTPVAIRCDGACSTATTE 378
+ P + A + VCTY E + +++LP C PGVDP ++PVA+ C CG C TT+
Sbjct: 64 VM--PAALPAIPQPVCTYRELRFASIRLPGCCPPGVDPMVSPVALSCHGPCQIKTTD 119

Figure 8

ClustalW alignment for NOVGN:

Based on this alignment, black outlined amino acid residues indicate regions of conserved sequence (i.e., regions that may be required to preserve structural or functional properties); greyed amino acid residues can be mutated to a residue with comparable steric and/or chemical properties without altering protein structure or function (e.g. L to V, I, or M); non-highlighted amino acid residues can potentially be mutated to a much broader extent without altering structure or function.

GTHB_CYPCA	M G T P V K T	L V V R N H I	L F S V V V L L	A V A Q S S	--- Y L P P C	E P V N E T V A V E K E G C P K C L V	L Q T T I
Q98849_GTHB2_GOLDFISH	M G T P V K I	L V V ---	L F S V V V L L	A V A Q S S	--- Y L P P C	E P V N E T V A V E K E G C P K C L V	L Q T T I
LSHB_BOVIN	M E M F Q C I	L L W ---	L L L G V A G V	W A S R G P	--- L R P L C	Q P I N A T A A E K E A C P V C I T	F T T S I
LSHB_SHEEP	M E M L Q G L	L L W ---	L L L G V A G V	W A S R G P	--- L R P L C	Q P I N A T A A E K E A C P V C I T	F T T S I
novel_gonadotropin	M K L A F L F	L G P M A L L	L L A G Y G C V	L G A S S G N L R T F V G C	A V R E F T F L A K K P G C R G	- L R T T I D A	
GTHB_CYPCA	C S G H C L T	- K E P V Y K S	P E S T V Y Q	H V C T Y R D V R V E T V R L	P D C P P G V D P H I	I T Y P V A L S C D C S L	
Q98849_GTHB2_GOLDFISH	C S G H C L T	- K E P V Y K S	P E S T V Y Q	H V C T Y R D V R V E T V R L	P D C P P G V D P H I	I T Y P V A L S C D C S L	
LSHB_BOVIN	C A G Y P S	- M K R V L P V I L	P P M P Q	R V C T Y H E L R E A S V R L	P G C P P G V D P M V S E P V A L S C H C G P		
LSHB_SHEEP	C A G Y P S	- M K R V L P V I L	P P M P Q	R V C T Y H E L R E A S V R L	P G C P P G V D P M V S E P V A L S C H C G P		
novel_gonadotropin	C W G R E T W E	K P I L E P F	I E A H H R V C T Y N E T K Q V I V K	L P N C A P G V D F F	Y T Y P V A I R E D C G A		
GTHB_CYPCA	C T M D T S D	---	---	---	---	---	---
Q98849_GTHB2_GOLDFISH	C T M D T S D	---	---	---	---	---	---
LSHB_BOVIN	C R L S S I D	---	---	---	---	---	---
LSHB_SHEEP	C R L S S I D	---	---	---	---	---	---
novel_gonadotropin	C S I A T T E	L R L M P G E A A V A L G F W C	Q R R R Q S S E T T G T R W R	A A V R D K V S L	L K A V D G W N G L L G		
GTHB_CYPCA	---	---	---	---	---	---	---
Q98849_GTHB2_GOLDFISH	---	---	---	---	---	---	---
LSHB_BOVIN	---	---	---	---	---	---	---
LSHB_SHEEP	---	---	---	---	---	---	---
novel_gonadotropin	D P A S S Q G L S A S S C T P V F P L A F Q I D S A S G K V G N F S S K Q T F I F S S A E I T L G G T						

Figure 9A

NOVINTRA A cDNA sequence:

Regions 116708 to 116812 (1-105bp), 117121 to 117248 (106-233p), and 117529 to 117778 (234-483bp).

CACTGTCATACTGTTTCAGAATTAAATATGCAGACCAGAAGGCTCTATACACAAGAG
ATGGCCAGCTGCTGGTGGGAGATCCTGTTGCAGACAACCTGCTGTGCAGAGAAGATC
TGCATACTTCCTAACAGAGGCTTGGCCCGCACCAAGGTCCCCATTTTCCTGGGGATC
CAGGGAGGGAGCCGCTGCCTGGCATGTGTGGAGACAGAAGAGGGGCCTTCCCTACA
GCTGGAGCCATCCACCTTGCCCCCACAGGATGTGAACATTGAGGAACTGTACAAAG
GTGGTGAAGAGGCCACACGCTTCACCTTCTTCCAGAGCAGCTCAGGCTCCGCCTTCA
GGCTTGAGGCTGCTGCCTGGCCTGGCTGGTTCCTGTGTGGCCCGGCAGAGCCCCAGC
AGCCAGTACAGCTCACCAAGGAGAGTGAGCCCTCAGCCCGTACCAAGTTTTACTTTG
AACAGAGCTGGTAGGGAGACAGGAAACTGC

Figure 9B

NOVINTRA A Protein sequence:

LSYCFRIKYADQKALYTRDGQLLVGDPVADNCCAEEKICILPNRGLARTKVPIFLGIQGG
RCLACVETEEGPSLQLEPSTLPPQDVNIEELYKGGEATRFTFFQSSSGSAFRLEAAAWP
GWFLCGPAEPQQPVQLTKESEPSARTKFYFEQSW

Figure 10A

BlastN for NOVINTRA A:

gb:GENBANK-ID:AF072476|acc:AF072476 Equus caballus interleukin-1 receptor
antagonist secretory form (IL-1ra) gene, partial cds - Equus caballus,
221
bp.

Top Previous Match Next Match

Length = 221

Plus Strand HSPs:

Score = 321 (48.2 bits), Expect = 8.1e-08, P = 8.1e-08

Identities = 147/218 (67%), Positives = 147/218 (67%), Strand = Plus / Plus

Query: 253 AGGATGTGAACATTGAGGAACGTACAA-AGGTGGTGAAGAGGCCACACGCTTCACCTTC 311
AGG GT AACAT GA CTG CAA A G GA GAG CA CGCTTCACCTTC
Sbjct: 4 AGGCAGTTAACATCACTGACCTGAGCAAGAACAAG-GAGGAGAACAAGCGCTTCACCTTC 62

Query: 312 TTCCAGAGCAG-CTCAGGCTCCGCCTTCAGGCTT-GAGGCTGCTGCCTGGCCTGGCTGGT 369
TCC G CA C GGC CC CC CAG CTT GAG CTGC GCCTG CCTGGCTGGT
Sbjct: 63 ATCC-GCTCAAACAGTGGCCCCACCACCAG-CTTCGAGTCTGCCGCCTGCCCTGGCTGGT 120

Query: 370 TCCTGTGTGGCCCGGCAG-AGCCCCAGCAGCCAGTACAGC-TCACCAAGGAG---AGTGA 424
TCCT TG CG CAG AG C A C GCC GT CAGC TCACCAA AG A GA
Sbjct: 121 TCCTCTGCACGGCG-CAGGAGGCAGACCGGCCCGT-CAGCCTCACCAACAAGCCCAAAGA 178

Query: 425 GCCCT-CAGCCCGT-ACCAAGTTTTACTTTGAACAGAGCTGGTAG 467
G CCT CA GT ACCAAGTT TACTT A AG C GTAG
Sbjct: 179 GTCCTTCATG--GTCACCAAGTTCTACTTCCAGGAGGACCAGTAG 221

Figure 10B

BlastX for NOVINTRA A:

ptnr:TREMBLNEW-ACC:CAB59831 IL-1L1 PROTEIN - Mus musculus (Mouse), 155 aa.

Top Previous Match Next Match

Length = 155

Plus Strand HSPs:

Score = 309 (108.8 bits), Expect = 8.0e-27, P = 8.0e-27

Identities = 70/149 (46%), Positives = 93/149 (62%), Frame = +3

Query: 12 CFRIKYADQKALYTRDGQLLVGDPVADNCC-AEKICILPNRGLARTKVPILFLGIQGGSRC 188
CFR+K + K LY + QLL G A+ E+I ++PNR L + P+ LG+QGGSC

Sbjct: 8 CFRMKDSALKVLYLHNNQLLAGGLHAEKVIKGEIISVVPNRALDASLSPVILGVQGGSC 67

Query: 189 LACVETEEGPSLQLEPSTLPPQDVNIEELYKGGEATRFTFFQSSSGSAFRLEAAAWPGW 368
L+C TE+GP L+LEP VNI ELY G +E+ FTF++ G E+AA+PGW

Sbjct: 68 LSC-GTEKGPILKLEP-----VNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGW 119

Query: 369 FLCGPAEPQQPVQLTKESEPSAR----TKFYFEQ 458

FLC E QPV+LT+ E A T FYF+Q

Sbjct: 120 FLCTSPEADQPVRILTQIPEDPAWDAPITDFYFQQ 153

>ptnr:SPTREMBL-ACC:Q9UBH0 FIL1 DELTA (INTERLEUKIN-1 LIKE PROTEIN 1)
(INTERLEUKIN-1 RECEPTOR ANTAGONIST HOMOLOG 1) (INTERLEUKIN-1 DELTA) - Homo
sapiens (Human), 155 aa.

Plus Strand HSPs:

Score = 285 (100.3 bits), Expect = 4.3e-24, P = 4.3e-24

Identities = 67/149 (44%), Positives = 89/149 (59%), Frame = +3

Query: 12 CFRIKYADQKALYTRDGQLLVGDPVADNCC-AEKICILPNRGLARTKVPILFLGIQGGSRC 188
CFR+K + K LY + QLL G A E+I ++PNR L + P+ LG+QGGSC

Sbjct: 8 CFRMKDSALKVLYLHNNQLLAGGLHAGKVIKGEIISVVPNRWLDASLSPVILGVQGGSC 67

Query: 189 LACVETEEGPSLQLEPSTLPPQDVNIEELYKGGEATRFTFFQSSSGSAFRLEAAAWPGW 368
L+C +E P+L LEP VNI ELY G +E+ FTF++ G E+AA+PGW

Sbjct: 68 LSCGVGQE-PTLTLEP-----VNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGW 119

Query: 369 FLCGPAEPQQPVQLTKESEPSAR----TKFYFEQ 458

FLC E QPV+LT+ E T FYF+Q

Sbjct: 120 FLCTVPEADQPVRILTQLPENGGWNAPITDFYFQQ 153

Figure 11

ClustalW for NOVINTRA A:

Based on this alignment, black outlined amino acid residues indicate regions of conserved sequence (i.e., regions that may be required to preserve structural or functional properties); greyed amino acid residues can be mutated to a residue with comparable steric and/or chemical properties without altering protein structure or function (e.g. L to V, I, or M); non-highlighted amino acid residues can potentially be mutated to a much broader extent without altering structure or function.

Sequences used:

- 1) X86458spoil1 (cDNA encoding a human spoil-I protein); X86459spoil2 (cDNA encoding a human spoil-II protein); X89432-il1delta (human IL-1 delta encoding cDNA); Q14628 (intracellular IL-1 receptor antagonist type II-human); G512010il1beta (ovine IL-1 beta - pig)

```

X86458spoil1  - - - - - DNHTMRGTPGDA - - - - - DGG - - GRAVYQS - - - - -
X86459spoil2  - - - - - DNHTMRGTPGDA - - - - - DGG - - GRAVYQSSESNVGMGLWRLRPSALTLS PV
X89432-il1delta - - - - - MALADLYE EG - - - - - GGG - - GGEGEDNADSK - - - - - E - - - - -
Q14628        - - - - - MALADLYE EG - - - - - GGG - - GGEGEDNADSK - - - - - E - - - - -
il1beta-1DNA  - - - - - MALADLYE EG - - - - - GGG - - GGEGEDNADSK - - - - - E - - - - -
G512010il1beta MATVPEP I NEV MAYYSDE NELLFEVDGPKQMKSCTQHLDLGSMGDGNIQLQISHQLYNKS

X86458spoil1  - - - - - MCKPIITGTINDLNQQVWTLQGGNVLVAVPRSDS - - MTPVTVAVIT
X86459spoil2  EAPAFSAP ECTLPFPFVCKPIITGTINDLNQQVWTLQGGNVLVAVPRSDS - - MTPVTVAVIT
X89432-il1delta - - - - - MVLSGALCFRMKDSALKVLYLHNNQLLAGGLHA - - - - - GKVIKGEEI
Q14628        - - - - - TICRPSGRKSSKMQAFRIWDVNQKTFYLRNNQLVAGYLQG - - - - - PNVNLEEKI
il1beta-1DNA  - - - - - GPSALSYCFRIKYADQKALYTRDGQLLVGDPVA - - - - - DNCCAEEKI
G512010il1beta FRQVVSVIAMEKLRSRAYEHVFRDDDLRSILSFI FEEEPVIFETSSDELUCDAAVQSVK

X86458spoil1  CKYPEAL - - - - - EQGRGDPIYLGIONPEM
X86459spoil2  CKYPEAL - - - - - EQGRGDPIYLGIONPEM
X89432-il1delta SVVPNEW - - - - - LDASLSPVILGVGGGSQ
Q14628        DVVP - - - - - IEPHALFLGIHGKMK
il1beta-1DNA  CILPNRG - - - - - LARTKVPIELGIQGGSR
G512010il1beta CKLQDREQKSLVLDSPCVLKALHLLSQEMSREVVFCMSFVQGEERDNKIPVALGIRDKNL

X86458spoil1  CLYCEKVGEQPTLQLKE - - - - - QKIMDLYGQPEPVKPFIFYRAKTGRTSTLESVAFPD
X86459spoil2  CLYCEKVGEQPTLQLKE - - - - - QKIMDLYGQPEPVKPFIFYRAKTGRTSTLESVAFPD
X89432-il1delta CLSCG-VGQEP TLTLEP - - - - - VNIMELYLGAKESKSFIFYRDMGLTSSSFESAAYPG
Q14628        CLSCVKS GDETRLQLEA - - - - - VNITDLS ENRKQDKRFATIESDSGPITTSFESAACPG
il1beta-1DNA  CLACVETE EGPSLQLEPSTLPPQD VNIEELYKGGE EATRFTFFQSSSGSAFRLEAAAWPG
G512010il1beta YLSCVKKGDTPTLQLEE - - - - - VDPKVPKRNMEKRFVFKTEIKNTVEFESVLYPN

X86458spoil1  WFIASS - KRDQPIILTSELGKS - - - - - YNTAFELNIND -
X86459spoil2  WFIASS - KRDQPIILTSELGKS - - - - - YNTAFELNIND -
X89432-il1delta WFLCTVPEADQPVRLTQLPENGGWNAPI TDFYFQQCD -
Q14628        WFLCTAMEADQPVSLTNMPDEG - - - - - VMVTKFYFQED E -
il1beta-1DNA  WFLCGPAEPQQPVLTK ESEPS - - - - - ARTKFFYFEQSW -
G512010il1beta WYISTSQIE EKPVFLGRFRGGQ - - - - - DI TDFRMETLSP

```

Figure 12A

NOVINTRA B cDNA sequence:

Regions 116968 to 117028 (1-61bp), 104067 to 104164 (62-159bp), 101958 to 102098 (160-300bp), and 100915 to 101134 (301-520bp) of genomic clone (-frame).

**ATGGGCACACCTGGCCTGGCCCTGCATGCAGACTGGACGGTGAGCCAGGACTTCTG
CAGGACACCCAAATCCTATGCTATTCGTGATTCTCGACAGATGGTGTGGGTCCTGAG
TGGAAATTCTTTAATAGCAGCTCCTCTTAGCCGCAGCATTAAAGCCTGTCACTCTTCAT
TTAATAGCCTGTAGAGACACAGAATTCAGTGACAAGGAAAAGGGTAATATGGTTTA
CCTGGGAATCAAGGGAAAAGATCTCTGTCTCTTCTGTGCAGAAATTCAGGGCAAGCC
TACTTTGCAGCTTAAGGAAAAAAATATCATGGACCTGTATGTGGAGAAGAAAGCAC
AGAAGCCCTTTCTCTTTTCCACAATAAAGAAGGCTCCACTTCTGTCTTTCAGTCAGT
CTCTTACCCTGGCTGGTTCATAGCCACCTCCACCACATCAGGACAGCCCATCTTTCTC
ACCAAGGAGAGAGGCATAACTAATAACACTAACTTCTACTTAGATTCTGTGGAATA
AATCCAGC**

Figure 12B

NOVINTRA B Protein sequence:

**MGTPGLALHADWTVSQDFCRTPKSYAIRDSRQMVWVLSGNSLIAAPLSRSIKPVTLHLI
ACRDTEFSDKEKGNMVYLGIGKDLCLFCAEIQGKPTLQLKEKNIMDLVEKKAQKPFL
FFHNKEGSTSVFQSVSYPGWFIATSTTSGQPIFLTKERGITNNTNIFYLDSVE**

Figure 13A

BlastN for NOVINTRA B:

gb:GENBANK-ID:PIGIRAP|acc:L38849 Sus scrofa (clone p0328w).IRAP1 mRNA,
complete cds - Sus scrofa, 563 bp (RNA).

Top Previous Match Next Match

Length = 563

Plus Strand HSPs:

Score = 309 (46.4 bits), Expect = 1.1e-07, P = 1.1e-07

Identities = 197/331 (59%), Positives = 197/331 (59%), Strand = Plus / Plus

Query: 169 CATTTAATAGCCTGTA-GAGAC-ACAGAATTCAGTGACAAGGAAAAGGGTAATATGGTTT 226
CA TA TA CTG A GAGA A AGA T GTG C G AG T AT T GT T
Sbjct: 217 CAAATACTAACTGGAAGAGAAGATAGATGTG-GTGCCT--GTTGAGCCTCATTGTGT 273

Query: 227 ACCTGGGAATCAAGGGAAAAGATCTCTGTCTCTTCTGTGCAGAAATTCAGGGCAAGCC-T 285
CCTGGG ATC A GGA A CT TG CT T CTGTG AA T C GG A G T
Sbjct: 274 TCCTGGGGATCCATGGAGGGAAGCTGTGCCTGTCTGTGTC-AAGT-CTGGTGATGAGAT 331

Query: 286 -ACTTTGCAGCTTAAGGAAAAAATATCATGGACCTGTATGTGGAGAAGAAAGCAC-AGA 343
A T CAG T A G A AA ATCA GACCTG A GA AG AGCA A A
Sbjct: 332 GAAGCTCCAGTTGGACGCAGTTAACATCACAGACCTG-AGAAAGAACAGCGAGCAGGACA 390

Query: 344 AGCCCTTTCTCTTTTTCCACAATAA-AGAAGGCTCCACTTCTGTCTTTTCAGTCAGTCTCT 402
AGC CTT CTT TCC C A AG GGC CCAC C CTTT AGTCAG C C
Sbjct: 391 AGCGCTTCACCTTCATCCGCTCCGACAGT-GGCCCCACCACCAGCTTTGAGTCAGCCGCC 449

Query: 403 TACCCTGGCTGGTTCATAGCCACCTCCACCACAT-CAGGACAGCCCATCTTTCTCACCAA 461
T CCTGGCTGGTTC T CAC T CAC A A CAG CAGCC T CTCACCAA
Sbjct: 450 TGTCTGGCTGGTTCCTCTGCAC-TGCACTAGAGGCAGACCAGCCTGTTGGCCTCACCAA 508

Query: 462 GGAGAG-AG-GCATACT-AATAACACTAACTTCTACTT 497
G A GCA T AA CAC AA TTCTACTT
Sbjct: 509 CACGCCCAAAGCAGCCGTCAAGGTCACCAAGTTCTACTT 547

Figure 13B

BlastX for NOVINTRA B:

>ptnr:SPTREMBL-ACC:Q9UHA5 FIL1 ETA - Homo sapiens (Human), 157 aa.
Plus Strand HSPs:

Score = 777 (273.5 bits), Expect = 3.2e-76, P = 3.2e-76
Identities = 149/149 (100%), Positives = 149/149 (100%), Frame = +1

Query: 64 PKSYAIRDSRQMVVWVLSGNSLIAAPLSRSIKPVTLHLIACRDTEFSDKEKGNMVYLGIGK 243
PKSYAIRDSRQMVVWVLSGNSLIAAPLSRSIKPVTLHLIACRDTEFSDKEKGNMVYLGIGK
Sbjct: 9 PKSYAIRDSRQMVVWVLSGNSLIAAPLSRSIKPVTLHLIACRDTEFSDKEKGNMVYLGIGK 68

Query: 244 KDLCLFCAEIQGKPTLQLKEKNIMDLYVEKKAQKPFLFFHNKEGSTSVFQSVSYPGWFIA 423
KDLCLFCAEIQGKPTLQLKEKNIMDLYVEKKAQKPFLFFHNKEGSTSVFQSVSYPGWFIA
Sbjct: 69 KDLCLFCAEIQGKPTLQLKEKNIMDLYVEKKAQKPFLFFHNKEGSTSVFQSVSYPGWFIA 128

Query: 424 TSTTSGQPIFLTKEGKITNNTNFYLDSE 510
TSTTSGQPIFLTKEGKITNNTNFYLDSE
Sbjct: 129 TSTTSGQPIFLTKEGKITNNTNFYLDSE 157

>ptnr:SPTREMBL-ACC:Q9NZH7 INTERLEUKIN-1 HOMOLOG 2 - Homo sapiens (Human), 164 aa.

Plus Strand HSPs:

Score = 411 (144.7 bits), Expect = 1.9e-37, P = 1.9e-37
Identities = 80/85 (94%), Positives = 81/85 (95%), Frame = +1

Query: 64 PKSYAIRDSRQMVVWVLSGNSLIAAPLSRSIKPVTLHLIACRDTEFSDKEKGNMVYLGIGK 243
PKSYAIRDSRQMVVWVLSGNSLIAAPLSRSIKPVTLHLIACRDTEFSDKEKGNMVYLGIGK
Sbjct: 9 PKSYAIRDSRQMVVWVLSGNSLIAAPLSRSIKPVTLHLIACRDTEFSDKEKGNMVYLGIGK 68

Query: 244 KDLCLFCAEIQGKPTLQLKEKNIMD 318
KDLCLFCAEIQGKPTLQLK + D
Sbjct: 69 KDLCLFCAEIQGKPTLQLKLQGSQD 93

>ptnr:TREMBLNEW-ACC:AAF02757 INTERLEUKIN-1 RECEPTOR ANTAGONIST HOMOLOG
- Homo sapiens (Human), 155 aa.

Top Previous Match Next Match

Plus Strand HSPs:

Score = 197 (69.3 bits), Expect = 5.9e-15, P = 5.9e-15
Identities = 52/147 (35%), Positives = 76/147 (51%), Frame = +1

Query: 73 YAIRDSRQMVVWVLSGNSLIAAPLSRSIKPVTLHLIACRDTEFSDKEKGNMVYLGIGKDL 252
+ ++DS V L N L+A L K + I+ + D ++ LG++G
Sbjct: 9 FRMKDSALKVLYLHNNQLLAGGLHAG-KVIKGEEISVVPNRWLDASLSPVI-LGVQGSQ 66

Query: 253 CLFCAEIQGKPTLQLKEKNIMDLYVEKKAQKPFLFFHNKEGSTSVFQSVSYPGWFIATST 432
CL C Q +PTL L+ NIM+LY+ K K F F+ G TS F+S +YPGW+ T
Sbjct: 67 CLSCGVGQ-EPTLTLEPVNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGWFLCTVP 125

Query: 433 TSGQPIFLTKEGKITNNTNFYLDSE 510
+ QP+ LT+ E G N T+FY +
Sbjct: 126 EADQPVRLTQLPENGGWNAPITDFYQQCD 155

Figure 14

ClustalW for NOVINTRA B:

Based on this alignment, black outlined amino acid residues indicate regions of conserved sequence (*i.e.*, regions that may be required to preserve structural or functional properties); greyed amino acid residues can be mutated to a residue with comparable steric and/or chemical properties without altering protein structure or function (*e.g.* L to V, I, or M); non-highlighted amino acid residues can potentially be mutated to a much broader extent without altering structure or function.

Sequences used:

- 2) X86459spoil2 (cDNA encoding a human spoil-II protein; Q14628 (intracellular IL-1 receptor antagonist type II-human); AAF02757 (intracellular IL-1 receptor antagonist type II-human); G512010; (OVINE IL-1 BETA - Pig)

Multiple Alignment:

```

il1beta2DNA      .....-MGTPLGLALHADWTVSQD-----
X86459spoil2     .....DNHTMRGTPGDADGGGGRAVYQSSSEN
Q14628           .....MALADLYEEEGGGGGEGEDNADSK-
AAF02757         .....
G512010il1beta  MATVPEPINEVMAYYSDENELLFEVDGPKQMKSC TQHLDLGSMGDGNIQLQISHQLYNKS

il1beta2DNA      .....FCRT.....PKSYAIR.....
X86459spoil2     AVGMGLWRLRP SALTLS PVEAPAFSAPLCTLPFP PVCK-PIITGTIN.....
Q14628           .....TICRPSGRKSSK-MQAFRIW.....
AAF02757         .....MVLSG-ALCFRMK.....
G512010il1beta  FRQVVSVIVAMEKLRSRAYEHVFRDDDLRSILSFI FEEEPVIFETSSDELLCDAAVQSVK

il1beta2DNA      ....DSROMVWVLSGNSLIAPLSRS IKPVTLHLTACRDTEFS DK-EKGNMVYLG IKKD
X86459spoil2     ....DLNQQVWTLQGONLVAVPRSDSVTPVTAVITCKYPEALEEQ-GRGDP IYLG IQNPE
Q14628           ....DVNQKTEFYLRNNQLVAGYIQ--GPNVNLEEKIDVVP-----IEPHALELGIHGGK
AAF02757         ....DSALKVLYLHNNQLLAGGLH--AGKVIKGE EISMV PNRWLD-ASLSPVILGVQGS
G512010il1beta  CKLQDREOKSLVLDSPCVLKAHLHLLSQEMSREVVFCMSFVQGEERDNKI PVALGIRDKN

il1beta2DNA      LCLFCAEHQ GKPTLQLKEKNIMDLYVEKKAQKPF LFFHNKEGSTSVFQSVSYPGWFIATS
X86459spoil2     MCLYCEKVGEQPTLQLKEQKIMDLYGQPEPVKPF LFYRAKTGRSTLESVAEPDWFIASS
Q14628           MCLSCVKSGDETRLQLEAVNITDLSENRKQDKRFAFIRSDSGPTTSFESAACP GWFLCTA
AAF02757         QCLSCG-VGQEP TLTLEPVNIMELYLGAKESKSF TFYRRDMGLTS SFESAAYPGWFLCTV
G512010il1beta  LYLSVCVKKGDTPTLQLEEVDPKVYP-KRNMEKRFFVYKTEIKNTVEFESVLYPNWYISTS

il1beta2DNA      TTSGQPIFLTKER--G--ITNNTNFYLDSE-
X86459spoil2     KRD-QPIILTSEL--G--KSYNTAFELNIND-
Q14628           MEADQPVSLTNMPDEG--VMVTKFYFQED-
AAF02757         PEADQPVRLTQLPENGGWNAPI TDFYFQQCD-
G512010il1beta  QIEEKPVFLGRFR--G--GQDIIDFRMETLSP
  
```

Figure 15A

NOVINTRA C cDNA sequence:

Region 168700 to 168880 (1-181bp), 169967 to 170118 (182-331bp) and 170635 to 170694 (332-391bp) of the genomic clone (- frame).

GATATCAATCATCGGGTGTGGGTTCTTCAGGACCAGACGCTCATAGCAGTCCCGAGG
AAGGTGTTCCCACTACTATTGCCTTAATCTCATGCCGACATGTGGAGACCCTTGAG
AAAGACAGAGGGAACCCCATCTACCTGGGCCTGAATGGACTCAATCTCTGCCTGAT
GTGTGCTAAAGTCGGGGACCAGCCCACACTGCAGCTGAAGCTTCAGGAAAAGGATA
TAATGGATTTGTACAACCAACCCGAGCCTGTGAAGTCCTTTCTCTTCTACCACAGCC
AGAGTGGCAGGAACTCCACCTTCGAGTCTGTGGCTTTCCCTGGCTGGTTCATCGCTG
TCAGCTCTGAAGGAGGCTGTCCTCTCATCCTTACCCAAGAACTGGGGAAAG

Figure 15B

NOVINTRA C Protein sequence:

DINHRVWVLQDQTLIAVPRKVPVTIALISCRHVETLEKDRGNPIYLGLNGLNLCLMCAK
VGDQPTLQLKLQEKDIMDLYNQPEPVKSFLFYHSQSGRNSTFESVAFPGWFIAVSSEGGC
PLILTQELGK

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 16A

BlastN for NOVINTRA C:

gb:GENBANK-ID:PIGIRAP|acc:L38849 Sus scrofa (clone p0328w) IRAP1 mRNA,
complete cds - Sus scrofa, 563 bp (RNA).

Top Previous Match Next Match

Length = 563

Minus Strand HSPs:

Score = 274 (41.1 bits), Expect = 5.6e-06, P = 5.6e-06

Identities = 150/244 (61%), Positives = 150/244 (61%), Strand = Minus / Plus

Query: 258 TCTACCTGGGCCTGAATGGACTCAATCTCTGCCGTGATGTGTGCTAAAGTCGGG-GACCAG 200
T T CCTGGG T ATGGA AA CT TGCTG TGTG T AAGTC GG GA AG
Sbjct: 271 TGTTCTGGGGATCCATGGAGGGAAGCTGTGCCTGTCCTGTG-TCAAGTCTGGTGATGAG 329

Query: 199 CCCACACTGCAGCTGAAGCTTCAGGAAAAGGATATAATGGATTTGTACAACCAACC-CGA 141
A CT CAG TG A C CAG AA AT A GA TG A AA AAC CGA
Sbjct: 330 ATGAAGCTCCAGTTGGA-CG-CAGTTAAC----ATCACAGACCTG-AGAAAGAACAGCGA 382

Query: 140 GCCTGTGAAGTCCTTTCTCTTCTACCACAGCCAGAGTGGCAGGAAGTCCACCTTCGAGTC 81
GC G AAG CTT CTTC CC C C A AGTGGC A C CCA CTT GAGTC
Sbjct: 383 GCAGGACAAGCGCTTCACCTTCATCCGCTCCGACAGTGGCCCCACCACCAGCTTTGAGTC 442

Query: 80 TGTGGCTTTCCCTGGCTGGTTCATCGCTGT-CAGCTCTGAAGG-AGGCTGTCCTCTCATC 23
G GC T CCTGGCTGGTTC TC TG C GC CT AGG AG C CCT T C
Sbjct: 443 AGCCGCCTGTCTGGCTGGTTCCTC--TGCACTGCACTAGAGGCAGACCAGCCTGTTGGC 500

Query: 22 CTTACCCA 15
CT ACC A
Sbjct: 501 CTCACCAA 508

Figure 16B

BlastX for NOVINTRA C:

>ptnr:SPTREMBL-ACC:Q9UHA7 FIL1 EPSILON - Homo sapiens (Human), 158 aa.
Plus Strand HSPs:

Score = 646 (227.4 bits), Expect = 2.4e-62, P = 2.4e-62
Identities = 126/130 (96%), Positives = 127/130 (97%), Frame = +1

Query: 1 DINHRVWVLQDQTLIAVPRK--VFPVTIALISCRHVETLEKDRGNPIYLGLNGLNLCLMC 174
DINHRVWVLQDQTLIAVPRK + PVTIALISCRHVETLEKDRGNPIYLGLNGLNLCLMC
Sbjct: 17 DINHRVWVLQDQTLIAVPRKDRMSPVTIALISCRHVETLEKDRGNPIYLGLNGLNLCLMC 76

Query: 175 AKVGDQPTLQLKLQEKDIMDLYNQPEPVKSFLFYHSQSGRNSTFESVAFPGWFIASSEG 354
AKVGDQPTLQLK EKDIMDLYNQPEPVKSFLFYHSQSGRNSTFESVAFPGWFIASSEG
Sbjct: 77 AKVGDQPTLQLK--EKDIMDLYNQPEPVKSFLFYHSQSGRNSTFESVAFPGWFIASSEG 134

Query: 355 GCPLILTQELGK 390
GCPLILTQELGK
Sbjct: 135 GCPLILTQELGK 146

>ptnr:SPTREMBL-ACC:Q9NZH8 INTERLEUKIN-1 HOMOLOG 1 - Homo sapiens, 169 aa.

Plus Strand HSPs:

Score = 416 (146.4 bits), Expect = 5.7e-38, P = 5.7e-38
Identities = 82/130 (63%), Positives = 101/130 (77%), Frame = +1

Query: 1 DINHRVWVLQDQTLIAVPRK--VFPVTIALISCRHVETLEKDRGNPIYLGLNGLNLCLMC 174
D+N +VW LQ Q L+AVPR V PVT+A+I+C++ E LE+ RG+PIYLG+ +CL C
Sbjct: 29 DLNQQVWTLQGQNLVAVPRSDSVTPVTVAITCKYPEALEQGRGDPYILGIQNPEMCLYC 88

Query: 175 AKVGDQPTLQLKLQEKDIMDLYNQPEPVKSFLFYHSQSGRNSTFESVAFPGWFIASSEG 354
KVG+QPTLQLK E+ IMDLY QPEPVK FLFY +++GR ST ESVAFP WFIA SS+
Sbjct: 89 EKVGEQPTLQLK--EQKIMDLYGQPEPVKPFIFYRAKTGRTSTLESVAFPDWFIA-SSKR 145

Query: 355 GCPLILTQELGK 390
P+ILT ELGK
Sbjct: 146 DQPIILTSELGK 157

>ptnr:SWISSPROT-ACC:P25085 INTERLEUKIN-1 RECEPTOR ANTAGONIST PROTEIN
PRECURSOR (IL-1RA) (IL-1RN) (IRAP) - Mus musculus (Mouse), 178 aa.

Top Previous Match Next Match

Minus Strand HSPs:

Score = 173 (60.9 bits), Expect = 2.1e-12, P = 2.1e-12
Identities = 35/81 (43%), Positives = 50/81 (61%), Frame = -1

Query: 259 IYLGLNGLNLCLMCAKVGDQPTLQLKLQEKDIMDLYNQPEPVKSFLFYHSQSGRNSTFES 80
++LG++G LCL CAK GD ++L+L+E +I DL E K F F S+ G ++FES
Sbjct: 82 VFLGIHGGKLCLSCAKSGDD--IKLQLEEVNITDLSKNKEEDKRFTFIRSEKGPTTSFES 139

Query: 79 VAFPGWFIASSEGCGPLILT 17
A PGWF+ + E P+ LT
Sbjct: 140 AACPGWFLCTTLEADRPVSLT 160

Figure 17

ClustalW for NOVINTRA C:

Based on this alignment, black outlined amino acid residues indicate regions of conserved sequence (i.e., regions that may be required to preserve structural or functional properties); greyed amino acid residues can be mutated to a residue with comparable steric and/or chemical properties without altering protein structure or function (e.g. L to V, I, or M); non-highlighted amino acid residues can potentially be mutated to a much broader extent without altering structure or function.

Sequences used:

- 1) IL1X_MOUSE: Locus - IL1X_MOUSE; Accession - P25085; PID - g124313
- 2) IL1X_HORSE: Locus - IL1X_HORSE; Accession - O18999; PID - g6166230
- 3) AAF02757_HUMAN: Locus - AF186094_1; Accession - AAF02757; PID - g6049805

```

IL1X_MOUSE      MEICWGPYSHLISLLILFHSEAACRPSGKRPCCKMOAFRIWDTNQKTFYLRNNQLIAGY
IL1X_HORSE      MEIRRRSVRHLIS-LLFLFYSETACHPLGKRPCCKMOAFRIWDTNQKTFYMRNNQLVAGY
AAF02757_Human  -----MVLSGALC-----FRMKDSALKVLYLHNNQLLAGG
Novel_IL1       -----DINHRVWVLQDQTLIAVP

IL1X_MOUSE      LGGPNIKLE---EKIDMVP-----IDLHSVFLGHHGGKLCCLSCAKSGDDIKLQLEEVN-
IL1X_HORSE      LQESNTKLQ---EKIDVVP-----IEPDALFLGLHGRKLCLACVKSGDEIRFOLEAVN-
AAF02757_Human  LHAGKVIKG---EEISVVPNRWLDASLSPVILGVQGGGSQLSCG-VGQEPTLTLEPVN-
Novel_IL1       RKVFPVTIALISCRHYETLEK---DRGNPIYLGLNGLNLCLMCAKVGDQPTLQLKLEK

IL1X_MOUSE      -ITDLSKNKEEDKRFTFIRSEKGPTTSFESAACPGWFLCTTLEADRPVSLTNTPEE---P
IL1X_HORSE      -ITDLSKNKEENKRFTFIRSNKSGPTTSFESAACPGWFLCTAQEADRPVSLTNKPK---S
AAF02757_Human  -IMELYLGAKESKSFYFYRRDMGLTSSFESAAYPGWFLCTVPEADQPVRLTLQLPENGAWN
Novel_IL1       DIMDLYNQPEPVKSFLFYHSQSGRNSTFESVAFPGWFIAVSSSEGGCPILTLQELGK---

IL1X_MOUSE      LVTKFYFQEDQ
IL1X_HORSE      FMVTKFYLOEDQ
AAF02757_Human  APITDFYFQQCD
Novel_IL1       -----

```

Figure 18

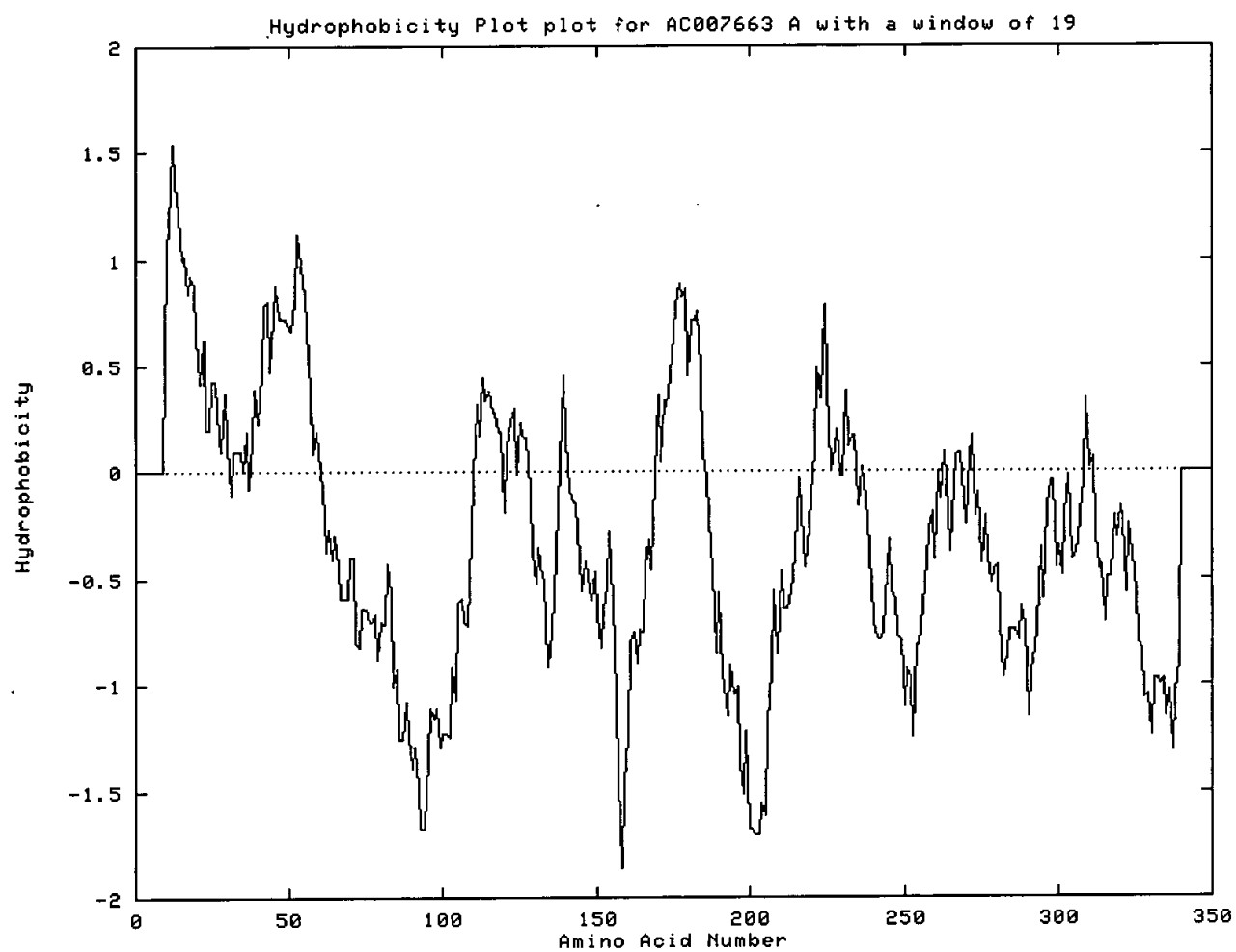
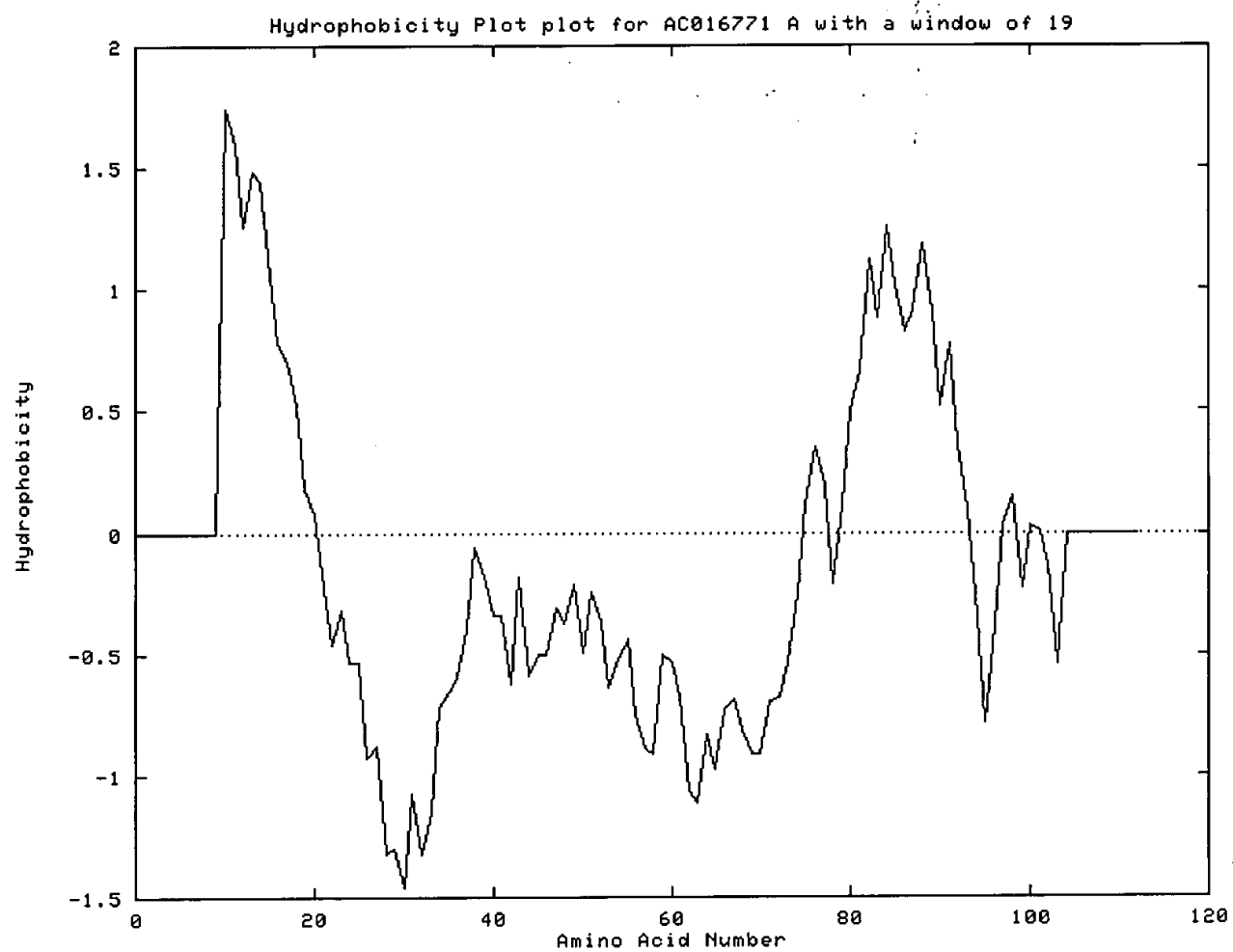


Figure 19



	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	2101	2102	2103	2104	2105	2106	2107	2108	2109	2110	2111	2112	2113	2114	2115	2116	2117	2118	2119	2120	2121	2122	2123	2124	2125	2126	2127	2128	2129	2130	2131	2132	2133	2134	2135	2136	2137	2138	2139	2140	2141	2142	2143	2144	2145	2146	2147	2148	2149	2150	2151	2152	2153	2154	2155	2156	2157	2158	2159	2160	2161	2162	2163	2164	2165	2166	2167	2168	2169	2170	2171	2172	2173	2174	2175	2176	2177	2178	2179	2180	2181	2182	2183	2184	2185	2186	2187	2188	2189	2190	2191	2192	2193	2194	2195	2196	2197	2198	2199	2200	2201	2202	2203	2204	2205	2206	2207	2208	2209	2210	2211	2212	2213	2214	2215	2216	2217	2218	2219	2220	2221	2222	2223	2224	2225	2226	2227	2228	2229	2230	2231	2232	2233	2234	2235	2236	2237	2238	2239	2240	2241	2242	2243	2244	2245	2246	2247	2248	2249	2250	2251	2252	2253	2254	2255	2256	2257	2258	2259	2260	2261	2262	2263	2264	2265	2266	2267	2268	2269	2270	2271	2272	2273	2274	2275	2276	2277	2278	2279	2280	2281	2282	2283	2284	2285	2286	2287	2288	2289	2290	2291	2292	2293	2294	2295	2296	2297	2298	2299	2300	2301	2302	2303	2304	2305	2306	2307	2308	2309	2310	2311	2312	2313	2314	2315	2316	2317	2318	2319	2320	2321	2322	2323	2324	2325	2326	2327	2328	2329	2330	2331	2332	2333	2334	2335	2336	2337	2338	2339	2340	2341	2342	2343	2344	2345	2346	2347	2348	2349	2350	2351	2352	2353	2354	2355	2356	2357	2358	2359	2360	2361	2362	2363	2364	2365	2366	2367	2368	2369	2370	2371	2372	2373	2374	2375	2376	2377	2378	2379	2380	2381	2382	2383	2384	2385	2386	2387	2388	2389	2390	2391	2392	2393	2394	2395	2396	2397	2398	2399	2400	2401	2402	2403	2404	2405	2406	2407	2408	2409	2410	2411	2412	2413	2414	2415	2416	2417	2418	2419	2420	2421	2422	2423	2424	2425	2426	2427	2428	2429	2430	2431	2432	2
--	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	---

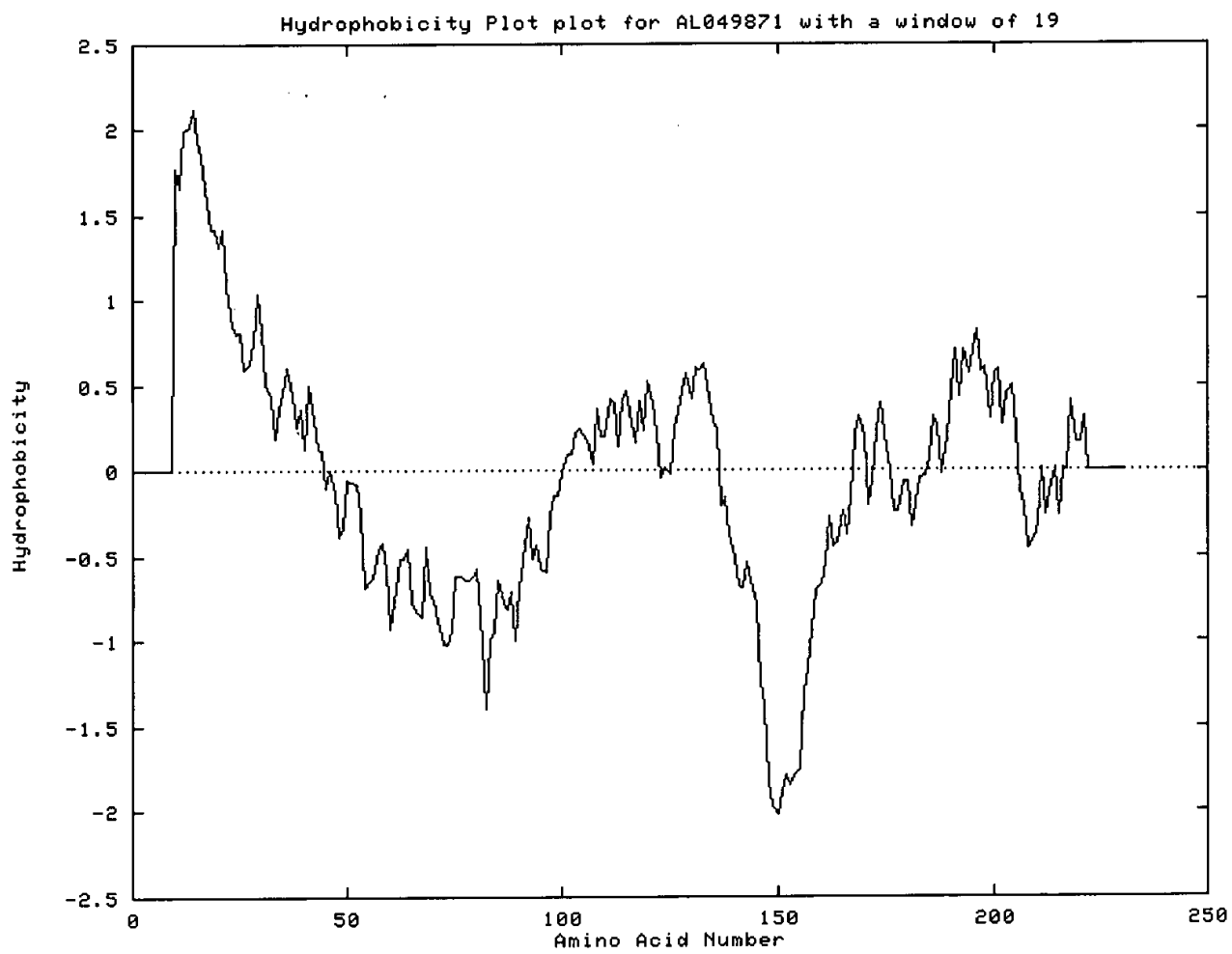


Figure 21

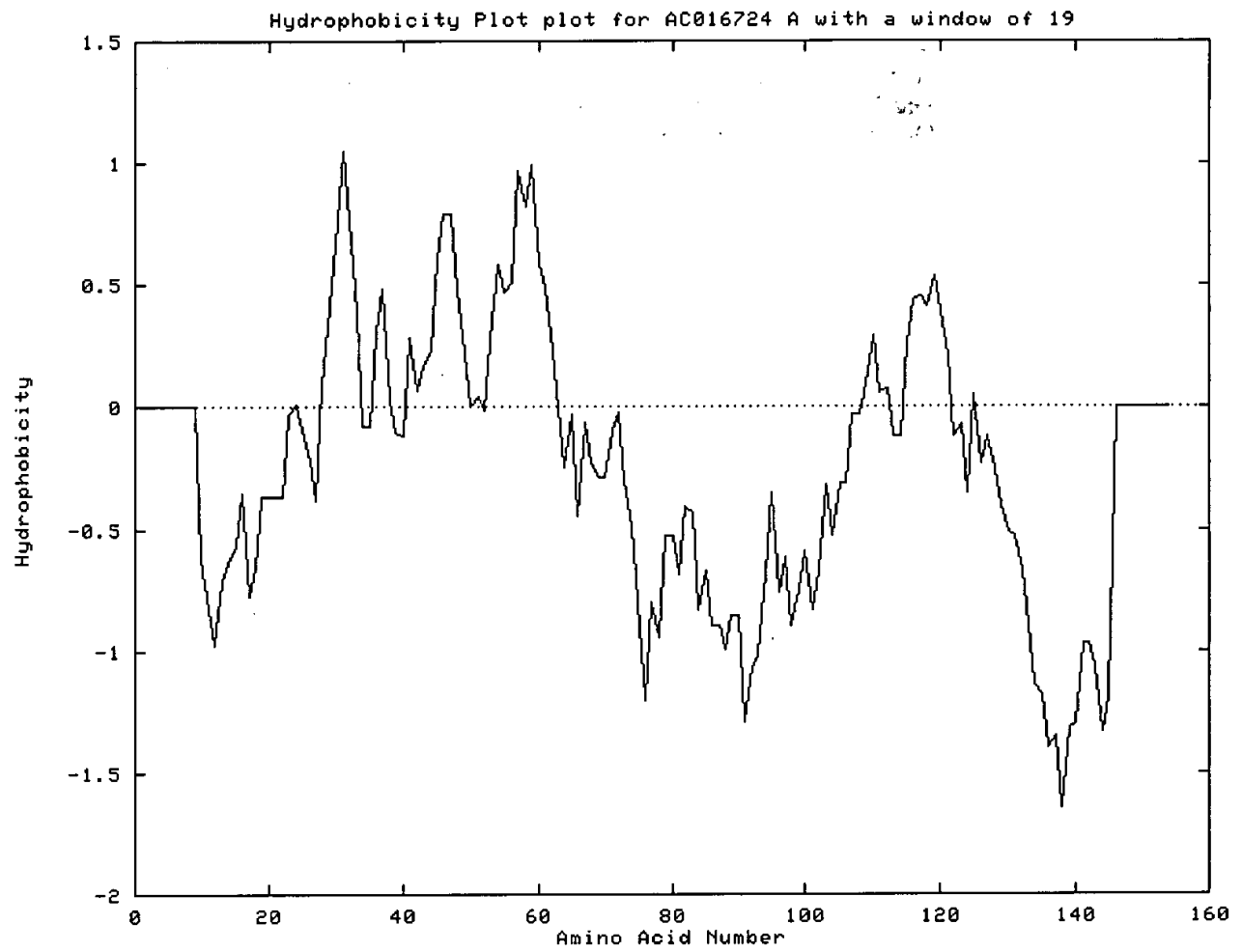


Figure 22

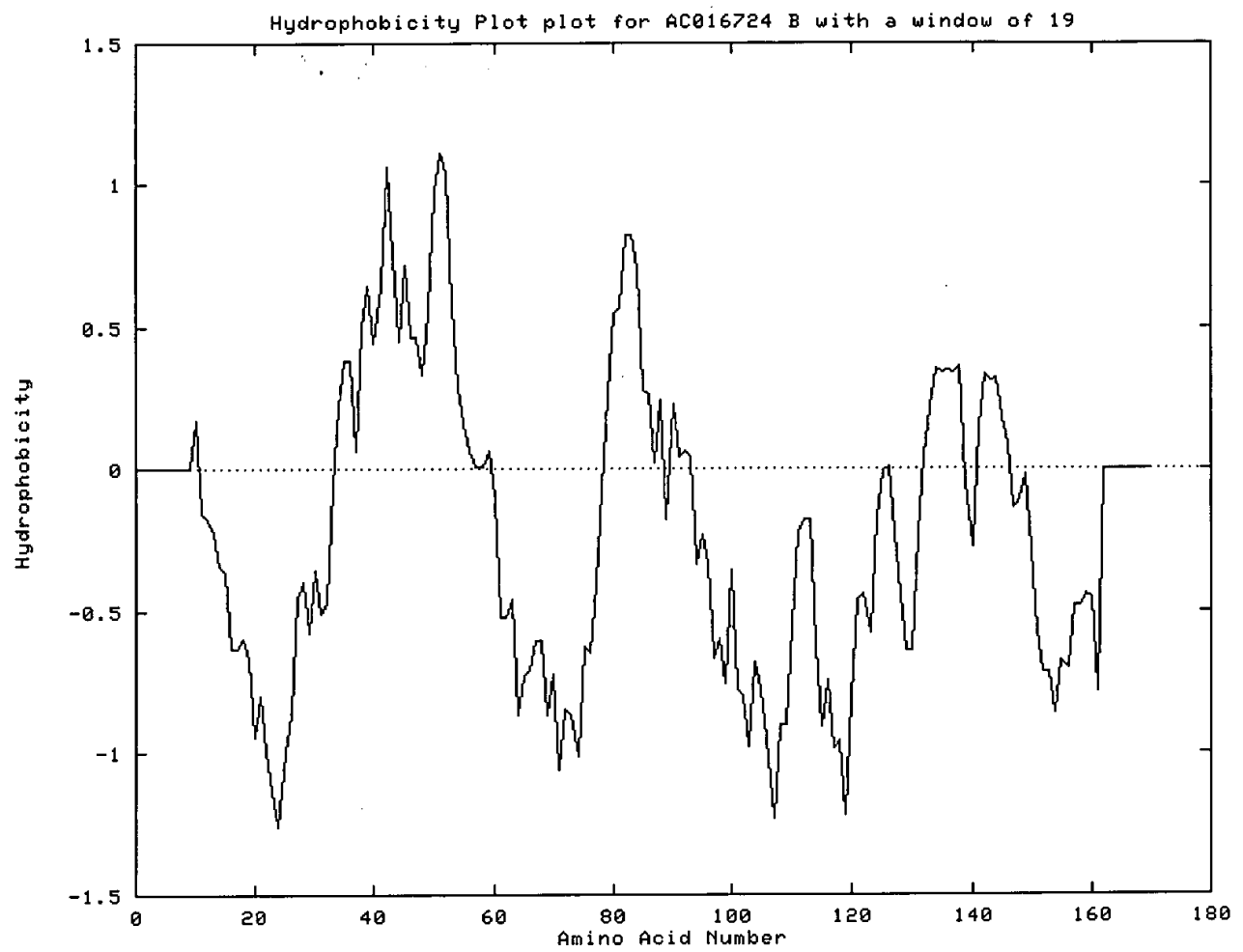


Figure 23

